

THE HEGSELMANN-KRAUSE DYNAMICS ON THE CIRCLE CONVERGE

PETER HEGARTY^{1,2}, ANDERS MARTINSSON^{1,2} AND EDVIN WEDIN^{1,2}

ABSTRACT. We consider the Hegselmann-Krause dynamics on a one-dimensional torus and provide the first proof of convergence of this system. The proof requires only fairly minor modifications of existing methods for proving convergence in Euclidean space.

1. INTRODUCTION

The so-called Hegselmann-Krause bounded confidence model (HK-model for brevity), introduced in [4], is one of the most popular mathematical models for the dynamics of opinion formation in groups of interacting agents. In its classical formulation, we have a finite number n of agents, indexed by the integers $1, 2, \dots, n$. Time is measured discretely and the opinion of agent i at time $t \in \mathbb{N} \cup \{0\}$ is represented by a real number $x_t(i)$. There is a fixed parameter $r > 0$ such that the dynamics are given by

$$(1.1) \quad x_{t+1}(i) = \frac{1}{|\mathcal{N}_t(i)|} \sum_{j \in \mathcal{N}_t(i)} x_t(j),$$

where $\mathcal{N}_t(i) = \{j : |x_t(j) - x_t(i)| \leq r\}$. In words, at each time step an agent takes account of his so-called r -neighbourhood, consisting of those other agents whose opinions currently lie within distance r of his own. He updates his opinion to the average of those held by the members of his r -neighbourhood, including himself.

In purely formal terms, the model makes sense if opinions $x_t(i)$ are assumed to come from a set V which has enough structure so that it is possible to make sense of the command to

‘move to the average of a finite collection of points within distance r of your present location’.

We shall refer to any rigorous formulation of this procedure as a *HK-update rule*. The simplest generalisation would be to \mathbb{R}^k for any $k \geq 1$ and with the Euclidean metric. This is natural if we imagine that there are k ‘issues’ on which agents have opinions, and they are only willing to compromise with those whose opinions on *every* issue are sufficiently close to their own. There are rigorous results in the literature concerning the HK-dynamics in Euclidean space of arbitrary dimension, see for example [1]. Note, though, that it is not clear what is the most natural formulation of the HK-dynamics in \mathbb{R}^k for $k > 1$. For example, the L^∞ -metric may better capture the notion that an agent will only compromise with those whose opinions on all issues are close to their own – indeed, the L^2 -metric seems better suited when considering the HK-dynamics as a rendezvous procedure, as in for example [5]. However, it is also plausible that the HK-update rule is too strict a criterion for compromise. At least one group of authors [10] has suggested a more general model in which there is a second parameter $l \leq k$ such that one agent will compromise with another whose opinions on at least l of the k issues are within a fixed distance of his own. Note that, in this generality, we are no longer dealing with a HK-update rule, as formulated above, since if $l < k$ then the ‘distance’ between agents does not satisfy the triangle inequality.

If we want to stick with the HK-update rule, then perhaps the simplest and most natural example to consider after Euclidean space is a circle. Let \mathcal{C}_p denote a circle of perimeter p ,

Date: April 7, 2015.

2000 Mathematics Subject Classification. 39A60, 93A14, 91D10.

Key words and phrases. Hegselmann-Krause model, circle, convergence.

equipped with its natural metric which we will denote by $\delta(\cdot, \cdot)$. If $p \leq 2r$ then the ‘average location of those in your r -neighbourhood’ is not well-defined, unless there is a priori agreement on an orientation of the circle and an origin, i.e.: on a continuous bijection from \mathcal{C}_p to $[0, p)$. In any case, if such a priori agreement existed, then the dynamics would become trivial as any configuration would collapse immediately to a complete consensus. The circle is also uninteresting if $p > nr$, since then there must be a pair of consecutive agents at distance greater than r apart, so the configuration evolves as if it were living on the real line, with these two agents as the extreme opinions. Hence, when studying the circle, we may assume that $2r < p \leq nr$. Indeed, without loss of generality we can fix $r = 1$ and have $p \in (2, n]$ as the only variable parameter.

There are at least two possible motivations for considering the HK-dynamics on a circle:

- (i) Firstly, it is possible to think of ‘real-life’ situations where it makes more sense to consider opinions as lying on a circle instead of the real line. For example, the subject under debate may be the choice of a time of day or date on which to hold some event.
- (ii) Secondly, the circle is perhaps the simplest example of a space on which the HK-dynamics behave differently from in Euclidean space in a fundamental sense which we now describe.

It is a well-known fact that, in Euclidean space, any configuration of opinions obeying the dynamics in (1.1) will freeze in finite time, that is, there will be some $T > 0$ such that $x_t(i) = x_T(i)$ for all $1 \leq i \leq n$ and $t \geq T$. In fact, it is known that the time taken to freeze is bounded above by a universal polynomial function of the number n of agents. In one dimension, the best bound to date is $O(n^3)$, proven independently in [1] and [7]. In higher dimensions, the best published bound [3] is $O(n^8)$, but a recent preprint [6] improves this to $O(n^4)$. Moreover, in a frozen configuration in \mathbb{R}^k , agents must either *agree*, that is $x(i) = x(j)$, or be beyond each other’s influence, that is $d(x(i), x(j)) > r$.

On the other hand, if on the circle we place $n \geq p$ agents at angles $2\pi j/n$, $0 \leq j < n$, then we have a frozen configuration in which no proper subset of the agents is isolated from its complement, but nor are any two agents in agreement. There also exist frozen configurations with these properties in which agents are not equally spaced. It is easy to see that any such configuration must include at least 4 agents, and that the only example with 4 agents, modulo translations, is that in Figure 1 below. Furthermore, on the circle there are configurations which never freeze. For example, if we perturb the configuration in Figure 1 by a sufficiently small amount then we will obtain one starting from which we never freeze, since the updated configurations will converge back to that in Figure 1, possibly translated, at an exponential rate. See Remark 2.4 below for a proof of this last statement.

Before proceeding, let us formally define the term *convergence* for a finite set of agents obeying the HK-dynamics on a circle:

Definition 1.1. Let $n \in \mathbb{N}$, $p \in \mathbb{R}_{>2}$. Let $\mathbf{x}_0 = (x_0(1), \dots, x_0(n)) \in (\mathcal{C}_p)^n$, the latter equipped with the product topology. Suppose that, for all $t > 0$, the vectors $\mathbf{x}_t = (x_t(1), \dots, x_t(n))$ are defined inductively by the HK-update rule, with $r = 1$. If, for some $\mathbf{x}_\infty \in (\mathcal{C}_p)^n$, we have $\mathbf{x}_t \rightarrow \mathbf{x}_\infty$ in the product topology, then the sequence of configurations \mathbf{x}_t is said to *converge* to \mathbf{x}_∞ . It is common to abuse terminology slightly and say that the initial configuration \mathbf{x}_0 converges to \mathbf{x}_∞ .

In [2] it is stated as an open problem for \mathcal{C}_p whether any initial configuration will converge. The only other explicit mention of this problem that we could find is in [9] and, as far as we know, it has remained open up to now. Our first main result gives a positive answer:

Theorem 1.2. *For any $n \in \mathbb{N}$, $p \in \mathbb{R}_{>2}$ and $\mathbf{x}_0 \in (\mathcal{C}_p)^n$, the sequence (\mathbf{x}_t) defined by the HK-update rule converges.*

The proof of this result will be given in the next section. It requires only moderate variations on ideas already presented in the literature when studying the HK-model in Euclidean space. Two important and well-known concepts are involved. Let $\mathbf{x} \in (\mathcal{C}_p)^n$. Firstly, we define the

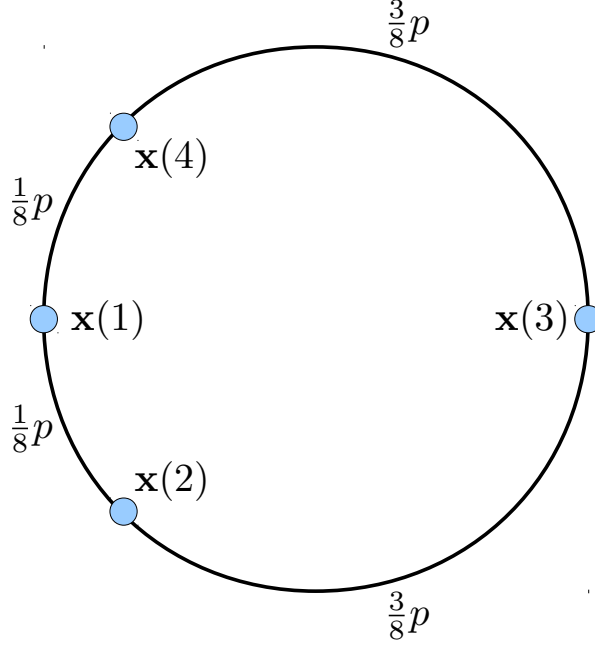


FIGURE 1. A frozen configuration in which agents are not equally spaced. The perimeter p should lie in the open interval $(2, \frac{8}{3})$.

energy of the configuration \mathbf{x} as

$$(1.2) \quad E(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n \min \{1, \delta(x(i), x(j))^2\}.$$

Secondly, we define the *influence graph* $G_{\mathbf{x}}$ corresponding to \mathbf{x} to be the undirected graph whose vertices are $1, 2, \dots, n$ and where an edge is placed between i and j if and only if $\delta(x(i), x(j)) \leq 1$. Given an orientation of the circle, we can also define a *directed influence graph* $\vec{G}_{\mathbf{x}}$ by choosing the oriented pair (i, j) if and only if the shortest path from i to j follows the orientation.

The proof of Theorem 1.2 consists of two main parts. Firstly, we show that an application of the HK-update rule always decreases the energy of a configuration, by an amount which is bounded below in terms of $\sum_{i=1}^n \delta(\mathbf{x}_t(i), \mathbf{x}_{t+1}(i))^2$. A similar result has been rediscovered several times over for Euclidean space, the earliest reference we could find is Theorem 2 of [8]. Our proof for the circle is almost identical to that for Euclidean space given in Theorem 4.3 of [1]. Secondly, we show that if the digraphs $\vec{G}_{\mathbf{x}_{t-1}}$ and $\vec{G}_{\mathbf{x}_t}$ are different, that is, if at least one edge has been either added, deleted or changed orientation at time t , then the sum of the decreases in energy at times t and $t+1$ is bounded below by $\frac{1}{9n^2}$. Since the initial energy is at most n^2 , this implies that the digraph $\vec{G}_{\mathbf{x}_t}$ can change at most $18n^4$ times and, in particular, must be fixed from some t onwards. Once we know this, convergence will follow from standard linear algebra arguments. Our reasoning in the second part of the proof also closely resembles known arguments for Euclidean space, though the latter all seem to appeal at some point to the fact that the points of \mathbf{x}_{t+1} all lie inside the convex hull of those in \mathbf{x}_t . There is no analogous fact for the circle, which gives the argument a new twist. The final proof of convergence, although it only appeals to basic linear algebra, also has no direct counterpart in Euclidean space since configurations always freeze eventually in the latter.

In Section 3, we will prove another result. Given that freezing occurs in Euclidean space after a time which is bounded by a universal polynomial function of the number n of agents and given that, on the circle, the directed influence graph changes at most a polynomial-in- n number of

times, it is natural to ask whether freezing of the latter must also happen within a time which is bounded by a polynomial function of n only. This turns out not to be the case. We shall prove that, for all sufficiently large n and any $T \in \mathbb{N}$, there is a choice of perimeter p and a configuration of n agents whose influence digraph remains the same up to time T but eventually changes. Indeed, we can choose the configuration so that $\vec{G}_{\mathbf{x}_t}$ is strongly connected all the way around the circle until the first change occurs, at which point a pair of adjacent agents become disconnected in the underlying graph and the configuration subsequently collapses to a complete consensus after three additional time steps.

Section 4 discusses some remaining open problems.

2. PROOF OF THEOREM 1.2

The first step in the proof is to show that the HK-dynamics decrease the energy of a configuration as defined in (1.2). The proof of the following result involves only minor modifications from that of Theorem 4.3 in [1], but we will present it for the sake of completeness.

Proposition 2.1. *If the sequence of configurations (\mathbf{x}_t) obeys the HK-update rule then, for any $t \geq 0$,*

$$(2.1) \quad E(\mathbf{x}_{t+1}) \leq E(\mathbf{x}_t) - 4 \cdot \sum_{i=1}^n \delta(\mathbf{x}_{t+1}(i), \mathbf{x}_t(i))^2.$$

Proof. Fix an (anti-clockwise) orientation and an origin of \mathcal{C}_p , i.e.: a suitable bijection $\phi : \mathcal{C}_p \rightarrow [0, p)$. In what follows we shall often not distinguish in writing between a point $x \in \mathcal{C}_p$ and its image $\phi(x)$ - hopefully, it will always be clear from the context which is being referred to.

For each $t \geq 0$, let $\Delta_t = (\Delta_t(1), \dots, \Delta_t(n)) \in [-1, 1]^n$ be the vector such that

$$(2.2) \quad \Delta_t(i) \equiv x_{t+1}(i) - x_t(i) \pmod{p}, \quad i = 1, \dots, n.$$

Note that (2.1) can be written in the form

$$(2.3) \quad E(\mathbf{x}_{t+1}) \leq E(\mathbf{x}_t) - 4 \cdot \|\Delta_t\|^2,$$

where $\|\cdot\|$ is the Euclidean norm on \mathbb{R}^n .

For every $1 \leq i, j \leq n$ set

$$(2.4) \quad r_t(i, j) := \begin{cases} \delta(\mathbf{x}_t(i), \mathbf{x}_t(j)), & \text{if the shortest path from } \mathbf{x}_t(i) \text{ to } \mathbf{x}_t(j) \text{ is anti-clockwise,} \\ -\delta(\mathbf{x}_t(i), \mathbf{x}_t(j)), & \text{otherwise.} \end{cases}$$

We write $i \sim_t j$ if $\delta(\mathbf{x}_t(i), \mathbf{x}_t(j)) \leq 1$. Define functions $\mathcal{E}_t : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ as follows:

$$(2.5) \quad \mathcal{E}_t(\mathbf{u}, \mathbf{v}) := \sum_{i=1}^n \sum_{j=1}^n \mathcal{E}_t(\mathbf{u}, \mathbf{v}, i, j),$$

where, if $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$, then

$$(2.6) \quad \mathcal{E}_t(\mathbf{u}, \mathbf{v}, i, j) := \begin{cases} (u_i - v_j - r_t(i, j))^2, & \text{if } i \neq j \text{ and } i \sim_t j, \\ 1, & \text{if } i \neq j \text{ and } i \not\sim_t j, \\ (u_i + v_j)^2, & \text{if } i = j. \end{cases}$$

Clearly the functions \mathcal{E}_t are symmetric, i.e.: $\mathcal{E}_t(\mathbf{u}, \mathbf{v}) = \mathcal{E}_t(\mathbf{v}, \mathbf{u})$, and

$$(2.7) \quad E(\mathbf{x}_t) = \mathcal{E}_t(\mathbf{0}, \mathbf{0}).$$

It is easy to see that the function $\mathbf{u} \mapsto \mathcal{E}_t(\mathbf{u}, \mathbf{0})$ is strictly convex on \mathbb{R}^n and attains its global minimum at $\mathbf{u} = \Delta_t$. Hence if we define $f_t : \mathbb{R}^2 \rightarrow \mathbb{R}$ by $f_t(z, w) = \mathcal{E}_t(z\Delta_t, w\Delta_t)$ then

$$(2.8) \quad 1 = \operatorname{argmin}_z f_t(z, 0) = \operatorname{argmin}_w f_t(0, w).$$

It is also clear that f_t is a convex second-degree polynomial in z and w . Thus it must have the form

$$(2.9) \quad f_t(z, w) = A(z-1)^2 + A(w-1)^2 + Bzw + C, \quad \text{for some } A \geq 0, -2A \leq B \leq 2A.$$

In particular, $f(1, 1) \leq f(0, 0)$ which means that

$$(2.10) \quad \mathcal{E}_t(\Delta_t, \Delta_t) \leq \mathcal{E}_t(\mathbf{0}, \mathbf{0}).$$

Secondly, we claim that

$$(2.11) \quad \mathcal{E}_t(\Delta_t, \Delta_t) - \mathcal{E}_{t+1}(\mathbf{0}, \mathbf{0}) \geq 4 \cdot \|\Delta_t\|^2.$$

Note that (2.11), (2.10) and (2.7) would together imply (2.3). To prove (2.11) we simply note that, on the one hand, if $1 \leq i \leq n$ then

$$(2.12) \quad \mathcal{E}_t(\Delta_t, \Delta_t, i, i) = (2\Delta_t(i))^2, \quad \mathcal{E}_{t+1}(\mathbf{0}, \mathbf{0}, i, i) = 0,$$

while, if $i \neq j$, then it is easily verified that

$$(2.13) \quad |r_{t+1}(i, j)| \leq |r_t(i, j) - (\Delta_t(i) - \Delta_t(j))|$$

and hence

$$(2.14) \quad \mathcal{E}_t(\Delta_t, \Delta_t, i, j) \geq \mathcal{E}_{t+1}(\mathbf{0}, \mathbf{0}, i, j), \quad i \neq j.$$

Together (2.14) and (2.12) imply (2.11), so the proof of the proposition is complete. \blacksquare

The second step in the proof of Theorem 1.2 is to establish a lower bound on the decrease in energy resulting from any change in the directed influence graph.

Proposition 2.2. *Let $t > 0$ and suppose that $\vec{G}_{\mathbf{x}_{t-1}} \neq \vec{G}_{\mathbf{x}_t}$. Then, with notation as in (2.2),*

$$(2.15) \quad \|\Delta_t\|^2 + \|\Delta_{t+1}\|^2 \geq \frac{1}{9n^2}.$$

Proof. We again suppose that an anti-clockwise orientation of the circle has been fixed. The terms left and right will be used synonymously with clockwise and anti-clockwise respectively. If $\vec{G}_{\mathbf{x}_{t-1}} \neq \vec{G}_{\mathbf{x}_t}$ then a priori at least one of the following four things must have happened at time t :

- (i) some agent, i_1 say, has a new neighbour i_2 to his right. In other words, the oriented edge (i_1, i_2) is in $\vec{G}_{\mathbf{x}_t}$ but not in $\vec{G}_{\mathbf{x}_{t-1}}$. Note that it could still be the case (if $p \leq 3$) that $(i_2, i_1) \in \vec{G}_{\mathbf{x}_{t-1}}$, it will not matter for the argument we present below.
- (ii) some agent i_1 has a new neighbour i_2 to his left.
- (iii) some agent i_1 has lost a neighbour i_2 on his right.
- (iv) some agent i_1 has lost a neighbour i_2 on his left.

We suppose that (i) occurs - analogous arguments can be given in each of the other three cases. Pick an oriented edge (i_1, i_2) which is added at time t . If either of these agents moved a distance of at least $\frac{1}{3n}$ at step t , then $\|\Delta_t\|^2 \geq \frac{1}{9n^2}$ and we are done. Otherwise, the distance between these two agents at time t is still at least $1 - \frac{2}{3n}$. Now agent i_1 will draw agent i_2 to the left (i.e.: clockwise) at step $t + 1$. If agent i_2 were to move a distance of at least $\frac{1}{3n}$ to the left at step $t + 1$ then we would be done again. But since i_2 gained a neighbour to his left at time t and since the ordering of agents around the circle is always preserved by the HK-dynamics, he cannot have lost any neighbours to his left at time t . Hence, there are only two possible ways he can be prevented from moving at least $\frac{1}{3n}$ to his left at time $t + 1$:

- (a) there is some i_3 such that the oriented edge (i_2, i_3) is also added to the digraph at time t ,
or
- (b) the neighbours of i_2 at time $t - 1$, including himself, move on average sufficiently far to the right at time t to compensate for the appearance of i_1 . Let M be the total net movement to the right of these neighbours at time t . Then, since i_2 has certainly no more than n neighbours at time t and since i_1 is such a neighbour, his movement to the left at time $t + 1$ is bounded below by $\frac{1}{n} \left(1 - \frac{2}{3n} - M\right)$. Thus $M \geq \frac{2}{3} \left(1 - \frac{1}{n}\right)$. Since i_2 has at most $n - 1$ neighbours at time $t - 1$, by the Cauchy-Schwarz inequality the sum of their contributions to $\|\Delta_t\|^2$ is at least $\frac{M^2}{n-1}$, which is greater than $\frac{1}{9n^2}$ for all $n \geq 3$.

Thus we are done again unless (a) occurs. Now we can apply the same analysis to the pair (i_2, i_3) and so on. For the proposition to fail, we would have to have an infinite, anti-clockwise sequence of agents i_1, i_2, \dots such that each oriented edge (i_k, i_{k+1}) appeared at time t . But this would mean that, for every k , agent i_{k+1} moved further to the left than i_k at time t . Since there are only finitely many agents, this is impossible. ■

Proof. of Theorem 1.2. The energy of a configuration of n agents is certainly no more than n^2 . Hence, if the sequence of configurations (\mathbf{x}_t) obeys the HK-update rule, it follows from Propositions 2.1 and 2.2 that there are at most $18n^4$ values of t for which $\vec{G}_{\mathbf{x}_t} \neq \vec{G}_{\mathbf{x}_{t+1}}$. Thus it suffices to prove Theorem 1.2 under the assumption that the directed influence graph \vec{G} is the same for all time. Furthermore, we may assume that the underlying graph G is connected, as otherwise the dynamics are the same as for a configuration on the real line, in which case we know it will freeze in finite time.

Once again, let us fix a choice of a suitable bijection $\phi : \mathcal{C}_p \rightarrow [0, p)$. The map ϕ induces an ordering of the components of any $\mathbf{x} \in \mathcal{C}_p^n$. Let us assume that, for our initial configuration $\mathbf{x}_0 = (x_0(1), \dots, x_0(n))$, we have $\phi(x(1)) \leq \phi(x(2)) \leq \dots \leq \phi(x(n))$. Then, since the HK-dynamics will never cause agents to cross, for all t the agents $1, 2, \dots, n$ will retain an anti-clockwise ordering, even though their ordering with respect to ϕ may be cyclically shifted.

Let \mathcal{P}^n denote the set of probability vectors in \mathbb{R}^n , i.e.: $\mathcal{P}^n = \{(v_1, \dots, v_n)^T \in \mathbb{R}^n : v_i \geq 0, \sum_i v_i = 1\}$. Then there is a natural map $\mathbf{x}_t \mapsto \mathbf{x}_t^* \in \mathcal{P}^n$ defined by

$$(2.16) \quad \mathbf{x}_t^* = (x_t^*(1), \dots, x_t^*(n)) \text{ where } p \cdot x_t^*(i) = \begin{cases} \delta(x_t(i+1), x_t(i)), & \text{if } 1 \leq i \leq n-1, \\ \delta(x_t(1), x_t(n)), & \text{if } i = n. \end{cases}$$

Since the directed influence graph is fixed, there is a fixed $n \times n$ real matrix A such that the dynamics satisfy¹

$$(2.17) \quad \mathbf{x}_{t+1}^* = A \mathbf{x}_t^*, \quad \text{for all } t \geq 0.$$

By Proposition 2.1, we know that the differences $\mathbf{x}_{t+1}^* - \mathbf{x}_t^*$ tend to $\mathbf{0}$ as $t \rightarrow \infty$. Hence, letting $\mathbf{v} := (A - I_n)\mathbf{x}_0^*$, we have

$$(2.18) \quad \mathbf{x}_{t+1}^* - \mathbf{x}_t^* = A^t \mathbf{v} \rightarrow \mathbf{0} \text{ as } t \rightarrow \infty.$$

Hence, \mathbf{v} must be identically zero in those coordinates corresponding to each block in the Jordan normal form of A whose eigenvalue is at least one in absolute value. It follows that there exists a $c > 0$, only depending on A , such that $\|\mathbf{x}_{t+1}^* - \mathbf{x}_t^*\| = O(e^{-ct})$. Thus \mathbf{x}_t^* must converge to a limit, call it \mathbf{x}_∞^* , such that

$$(2.19) \quad \|\mathbf{x}_t^* - \mathbf{x}_\infty^*\| = O(e^{-ct}).$$

By the definition of energy in (1.2), it follows that there is some $E_\infty \in \mathbb{R}_{\geq 0}$ such that $E(\mathbf{x}_t) = E_\infty + O(e^{-\Omega(t)})$ and hence in turn, by Proposition 2.1, that $\|\Delta_t\|^2 = O(e^{-\Omega(t)})$. This establishes convergence of the sequence (\mathbf{x}_t) , in the sense of Definition 1.1. ■

Remark 2.3. An alternative argument for obtaining (2.19) is as follows. It is easy to see that the matrix A in (2.17) has non-negative entries. We claim that A is column-stochastic. To prove this, it suffices to prove that

$$(2.20) \quad \mathbf{1}^T A \mathbf{v} = \mathbf{1}^T \mathbf{v} \text{ for all } \mathbf{v} \in \mathbb{R}^n, \text{ where } \mathbf{1} = (1, \dots, 1)^T.$$

By definition, (2.20) holds for $\mathbf{v} = \mathbf{x}_t^*$ and any t . Now let \mathbf{v} be any probability vector. The transformation $\mathbf{v} \mapsto A\mathbf{v}$ describes an updating of a configuration of n agents, possibly involving a cyclic shift of the agents' indices. This may not coincide with the HK-update rule for this configuration as there is no guarantee that agents interact precisely with those within unit distance of themselves. However, exactly the same pairs of agents interact as would be the case if the configuration were some \mathbf{x}_t . In particular, this means that the anti-clockwise ordering of

¹Note that it is crucial that the directed influence graph be fixed for the same to be true of the transition matrix, it would not suffice in general for the underlying graph to be fixed.

agents will be preserved (no agents will cross), which implies that $A\mathbf{v}$ will also be a probability vector. Thus (2.20) holds for any probability vector, and hence for any $\mathbf{v} \in \mathbb{R}^n$.

Now A may not be regular, since there may be pairs of agents that have exactly the same neighbours in \vec{G} at $t = 0$. However, any such pair will immediately reach consensus at $t = 1$. If we start at $t = 1$ and replace the vector \mathbf{x}_t^* of (2.16) by the (possibly shorter) vector of non-zero distances, then it is easy to see that we have an analogue of (2.17) in which the transition matrix is now regular. Thus we will have exponential convergence to the steady-state, which yields (2.19).

Remark 2.4. With notation as above we have, for the configuration $\mathbf{x} \in (\mathcal{C}_p)^4$, $p \in (2, \frac{8}{3})$, of Figure 1 that

$$(2.21) \quad \mathbf{x}^* = \begin{bmatrix} 1/8 \\ 3/8 \\ 3/8 \\ 1/8 \end{bmatrix}, \quad \mathbf{x}^* = A\mathbf{x}^* \text{ where } A = \begin{bmatrix} 1/6 & 1/4 & 0 & 1/12 \\ 1/2 & 5/12 & 1/3 & 1/4 \\ 1/4 & 1/3 & 5/12 & 1/2 \\ 1/12 & 0 & 1/4 & 1/6 \end{bmatrix}.$$

The second eigenvalue of A is $\lambda_2 = \frac{1}{3}$ with eigenvector $\mathbf{v}_2 = (1, 1, -1, -1)^T$. If we choose \mathbf{x}_0 such that $\mathbf{x}_0^* = \mathbf{x}^* + \varepsilon \mathbf{v}_2$ then, for sufficiently small but non-zero ε (how small depends on how close p is to 2 or to $8/3$), the HK-updates will satisfy (2.17) and the configurations \mathbf{x}_t will converge to some translation of \mathbf{x}_0 without ever actually reaching it.

3. INFLUENCE GRAPHS WHICH TAKE ARBITRARILY LONG TO FREEZE

In this section we will prove that the time taken for the (directed or undirected) influence graph to freeze is unbounded as a function of n . To simplify some notation, we will assume the perimeter p to be fixed and instead allow r to vary in the interval $(\frac{p}{n}, \frac{p}{2})$.

Theorem 3.1. *For any sufficiently large n there is an initial configuration with n agents such that the time it takes for the influence graph to freeze can be made arbitrarily large by picking r appropriately. Hence, the time until the influence graph freezes is unbounded solely as a function of n .*

Proof. Pick positive integers m_1, m_2 such that $m_1 < m_2 < 2m_1$ and $n = 3m_1 + 2m_2$. This is possible provided n is sufficiently large².

Consider a state consisting of five clusters³ of sizes m_1, m_2, m_1, m_1, m_2 as illustrated in Figure 2. If $d_2 = (\frac{m_2}{m_1}) d_1$ then such a configuration will be stable for any choice of $r \in (d_2, 2d_1)$. Let \mathbf{x} denote such a stable state and for $y_0(1), y_0(2)$ sufficiently small positive numbers, let \mathbf{x}_0 denote the perturbation of \mathbf{x} as indicated in the figure.

It is easy to see that, as long as every cluster in \mathbf{x}_t sees exactly one other cluster on either side, the state at time $t + 1$ is the perturbation of \mathbf{x} by $y_{t+1}(1)$ and $y_{t+1}(2)$, defined recursively by

$$(3.1) \quad y_{t+1}(1) = \frac{m_2 y_t(2)}{2m_1 + m_2}$$

$$(3.2) \quad y_{t+1}(2) = \frac{m_1 y_t(1) + m_2 y_t(2)}{2m_1 + m_2}.$$

Note that these sequences are positive, bounded by $\max\{y_0(1), y_0(2)\}$ and tending to 0 as $t \rightarrow \infty$. In particular, this means that the distance between the adjacent clusters of equal size in \mathbf{x}_t is strictly less than, but tending to d_2 . We assume $y_0(1)$ and $y_0(2)$ are sufficiently small such that the distance between any other pairs of clusters remains bounded away from d_2 .

To complete the proof of the theorem, consider the HK-dynamics with initial state \mathbf{x}_0 as described above for r slightly smaller than d_2 . In this case, the influence graph will be constant until the distance between the clusters of equal size is greater than r , at which time the model

²Indeed, there is a solution for any $n \geq 12$ except if $n \in \{13, 14, 15, 16, 18, 20, 21, 23, 25, 28, 30, 35\}$.

³By a *cluster* we mean a set of agents in agreement.

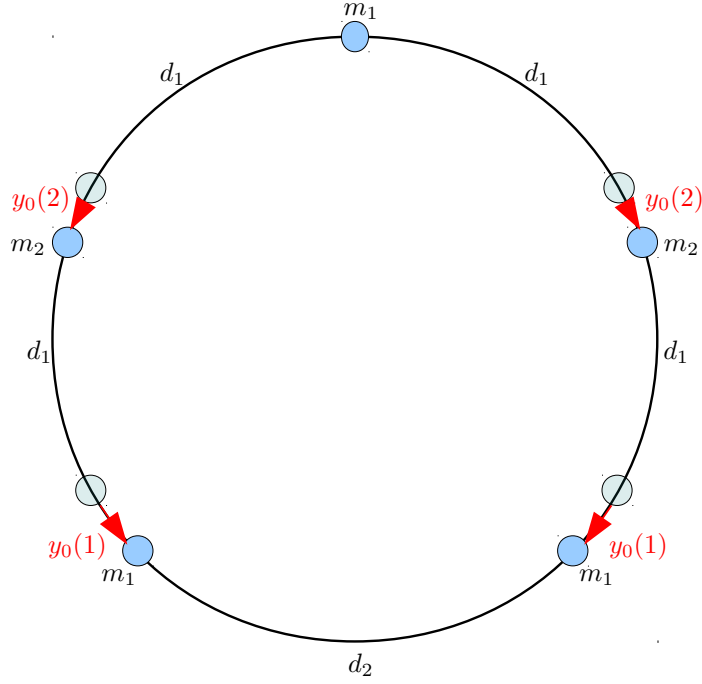


FIGURE 2. A state consisting of n agents in 5 clusters. We assume that every cluster sees one other cluster on each side. The inter-cluster distances d_1, d_2 refer to the stable state. The arrows indicate the perturbation from the stable state we consider.

behaves as five clusters on \mathbb{R} . The time before this happens can be made arbitrarily long by choosing r sufficiently close to d_2 . \blacksquare

Remark 3.2. Suppose we run the HK-dynamics on the real line, starting from five equally spaced clusters of sizes m_1, m_2, m_1, m_2 and m_1 , reading from left to right, with inter-cluster spacing d_1 and $r \in (d_1, d_2)$, where $d_2 = \left(\frac{m_2}{m_1}\right) d_1$ and $m_1 < m_2 < 2m_1$. It is easy to see that any such configuration will eventually reach a consensus. Hence, on the circle, the same will be true of the configuration in Theorem 3.1 for r sufficiently close to d_2 . We now show that the various parameters can be chosen so that, on the real line, consensus occurs after three time steps and hence, on the circle, three time steps after the graph disconnects.

First consider the real line again. At $t = 1$, the two extreme clusters will move inwards a distance $\left(\frac{m_2}{m_1+m_2}\right) d_1$ while the other three clusters will remain fixed. Hence if $\left(1 + \frac{m_1}{m_1+m_2}\right) d_1 < r < d_2$, which can be satisfied provided $m_2 > \sqrt{2}m_1$, then the middle cluster of size m_1 will be visible to all the others at $t = 1$. Hence at $t = 2$ we will have clusters of size $m_1 + m_2$ on either side of the middle cluster, and each at a distance of $\left(\frac{m_1^2+m_2^2+m_1m_2}{m_2(2m_1+m_2)}\right) d_1$ from it. Thus if

$$(3.3) \quad \frac{r}{d_1} \in \left(\frac{2(m_1^2 + m_2^2 + m_1m_2)}{m_2(2m_1 + m_2)}, \frac{m_2}{m_1} \right),$$

then the configuration will collapse to a complete consensus at $t = 3$. One may check that the open interval in (3.3) is non-empty if $\xi_0 < \frac{m_2}{m_1} < 2$, where $\xi_0 \approx 1.7693\dots$ is the unique real root of the equation $\xi^3 - 2\xi - 2 = 0$. It follows that also on a circle of fixed perimeter, for all n

sufficiently large and all $T \in \mathbb{N}$, there is a choice of r such that the configuration in Theorem 3.1 will remain connected up to time T and, once it disconnects, will collapse to consensus after three further time steps.

4. OPEN PROBLEMS

It is natural to ask whether the methods of this paper can be adapted to prove convergence of the HK-dynamics in even more general spaces than the circle. The case of higher-dimensional tori can be motivated in the usual manner, by imagining that there is more than one issue on which agents have opinions. Generalisation to even more abstract spaces would at present seem to be a purely mathematical curiosity. One of the main challenges in such work would be to give a mathematically rigorous formulation of the problem.

On the circle, there remain some questions of possible interest. One could investigate how the time to freezing of the influence digraph behaves for fixed p and r . This would seem to depend very subtly on either parameter. Another question that occurred to us but which we cannot presently resolve is whether the sequence of influence digraphs $\vec{G}_{\mathbf{x}_t}$ resulting from a given initial configuration must always be acyclic - in other words, if $\vec{G}_{\mathbf{x}_s} = \vec{G}_{\mathbf{x}_t}$ for some $s < t$, must it be the case that $\vec{G}_{\mathbf{x}_u} = \vec{G}_{\mathbf{x}_v}$ for all $s \leq u, v \leq t$?

5. ACKNOWLEDGEMENTS

We thank Bernadette Charron-Bost for pointing out a small error in an earlier version of the manuscript, in which we incorrectly stated that the matrix A in (2.17) was regular. Remark 2.3 presents the corrected assertions.

REFERENCES

- [1] A. Bhattacharya, M. Braverman, B. Chazelle and H. L. Nguyen, *On the convergence of the Hegselmann-Krause system*, Proceedings of the 4th Innovations in Theoretical Computer Science conference (ICTS 2013), Berkeley CA, January 2013.
- [2] V. D. Blondel, J. M. Hendrickx and J. N. Tsitsiklis, *On Krause's multi-agent consensus model with state-dependent connectivity*, IEEE Trans. Automat. Control 54 (2009), No. 11, 2586–2597.
- [3] S. R. Etesami and T. Basar, *Game-theoretic analysis of the Hegselmann-Krause model for opinion dynamics*, to appear in IEEE Trans. Autom. Control. Available online at <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7024142>
- [4] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence: models, analysis and simulations*, Journal of Artificial Societies and Social Simulation 5 (2002), No. 3. Fulltext at <http://jasss.soc.surrey.ac.uk/5/3/2/2.pdf>
- [5] S. Martinez, F. Bullo, J. Cortes and E. Frazzoli, *On Synchronous Robotic Networks - Part II: Time Complexity of Rendezvous and Deployment Algorithms*, IEEE Trans. Automat. Control 52 (2007), No. 12, 2214–2226.
- [6] A. Martinsson, *An improved energy argument for the Hegselmann-Krause model*. Preprint at <http://arxiv.org/pdf/1501.02183.pdf>
- [7] S. Mohajer and B. Touri, *On convergence rate of scalar Hegselmann-Krause dynamics*. Preprint at <http://arxiv.org/pdf/1211.4189v1.pdf>
- [8] M. Roozbehani, A. Megretski and E. Frazzoli, *Lyapunov analysis of quadratically symmetric neighborhood consensus algorithms*, Proceedings of the 47th IEEE Conference on Decision and Control (CDC 2008), pp. 2252–2257.
- [9] M. Zahri, *Condensing on metric spaces: modeling, analysis and simulation*, Ph.D. thesis, Goethe University Frankfurt (2009).
- [10] S. Zhang and Z. Sun, *Consensus analysis of a 2-dimensional KH model with various confidence level sets*, Proceedings of the Second International Conference on Intelligent Computation Technology and Automation (ICICTA 2009), IEEE Computer Society, Washington DC (2009).

¹MATHEMATICAL SCIENCES, CHALMERS, 41296 GOTHENBURG, SWEDEN

²MATHEMATICAL SCIENCES, UNIVERSITY OF GOTHENBURG, 41296 GOTHENBURG, SWEDEN

E-mail address: hegarty@chalmers.se

E-mail address: andemar@chalmers.se

E-mail address: edvinw@student.chalmers.se